Cell, Vol. 103, 9–12, September 29, 2000, Copyright ©2000 by Cell Press

Noncoding RNA Genes in Dosage Compensation and Imprinting

Minireview

Richard L. Kelley*‡ and Mitzi I. Kuroda*†‡
*Baylor College of Medicine
Department of Molecular and Cellular Biology
†Howard Hughes Medical Institute
Houston, Texas 77030

Students were once taught that there are three kinds of RNA: messenger RNA, ribosomal RNA, and transfer RNA, all required for protein synthesis. Today the list is far more extensive; for example, one RNA/protein complex acts on a second RNA substrate in splicing, editing, and posttranscriptional modifications, and ancient ribonucleoprotein complexes insert proteins into membranes and add telomeres onto the ends of chromosomes. To this growing list, we must add another class of RNAs, those associated with chromatin. Most RNA molecules appear to be soluble, so once transcription is complete the RNA is free of the chromosome. However, chromosomal RNAs act in *cis* near their sites of synthesis, suggesting a novel mechanism of RNA localization. Here we will focus on four unusual genes involved in dosage compensation which may act through RNA products: *Xist* in mouse, the antisense *Tsix* gene that downregulates *Xist*, and *roX1* and *roX2* in flies. The two mammalian genes may also reveal a link between RNA and imprinting.

Dosage Compensation

Diploid animals sometimes carry only a single copy of a particular chromosome due to mistakes made during meiotic segregation. This is almost always fatal during early embryonic development due to gross imbalances of hundreds of gene products. In contrast, males of many species are monosomic for the X, and yet survive through the action of dosage compensation which somehow equalizes X gene expression between males and females. Dosage compensation has evolved independently in many different animal groups using entirely distinct strategies (reviewed in Meller, 2000). Here we will discuss new findings about how dosage compensation is controlled in the mouse and point out unexpected parallels with the situation in *Drosophila*.

Xist *RNA in Mammals*

Mammals transcribe only one X chromosome per cell. Any additional X chromosomes (such as in normal XX:AA females) are silenced through the action of the X inactivation center (Xic) which includes the enigmatic *Xist* gene (*X i*nactive *s*pecific *t*ranscript). The *Xist* gene maps to the X chromosome and encodes a large (15–17 kb) spliced, polyadenylated RNA. It does not encode a protein, but instead has the remarkable ability to spread in *cis* from its site of synthesis to coat the inactive X. *Xist* is required to establish chromosome silencing, but is not essential for epigenetic maintenance of the silent state (see Meller, 2000 and references therein). The inactive X is marked by histone hypoacetylation, hypermethylation of CpG islands, and late replication, all hallmarks of heterochromatin. As silencing is established, females must somehow choose between two identical X homologs. How is a single X always protected from *Xist*-mediated inactivation? There are really two answers to this question depending upon when and where in the embryo one looks.

In the mouse, X inactivation can be separated into an imprinted pathway operating in the extraembryonic tissues, and a random choosing pathway in the embryo proper (reviewed in Solter and Wei, 1997). Imprinted genes are transcribed predominantly from just the maternal or paternal homolog (reviewed in Tilghman, 1999). In the case of *Xist*, the mother's X is imprinted to always remain transcriptionally active (*Xist* off). The father's X is somehow imprinted to be silenced (*Xist* on). These imprints persist throughout the preimplantation embryo. Around the time of implantation, lineages that form cells of the extraembryonic tissues retain the imprint and undergo preferential paternal silencing, while the epiblast lineage giving rise to the embryo proper erases the imprint. These cells subsequently count the number of X chromosomes, randomly choose one to remain active, and silence any others through the action of *Xist*.

Embryos can be created by nuclear transplantation which carry either two maternal or two paternal sets of chromosomes. Such embryos die due to a host of problems arising from imprinted genes. In addition, close examination of the extraembryonic tissues shows that they suffer incorrect dosage compensation in either case. If two X chromosomes are paternally inherited, both are silenced. If two are maternal, neither is silenced. The conclusion is that X chromosomes are not counted in the extraembryonic tissues as they are in the epiblast. Rather, the paternal X is imprinted to always be silenced, and the maternal X is somehow shielded from silencing.

Work from J. Lee (2000), in this issue of *Cell*, provides a molecular foundation for understanding the imprinting pathway by uncovering a role for *Tsix*. This novel gene occupies an overlapping DNA sequence with *Xist*, but is transcribed in the antisense direction (Figure 1A). *Tsix* was postulated to be a negative regulator of *Xist* based on the observation that when the *Tsix* promoter was destroyed, the mutant X was preferentially inactivated in female ES cells (Lee and Lu, 1999). In other words, when the antisense *Tsix* gene is not transcribed, *Xist* RNA is constitutively expressed in *cis*. In the new work, knockout mice are used to demonstrate that *Tsix* is also a prime candidate for an imprinting factor.

The *Tsix* promoter knockout shows a dramatic parent-of-origin phenotype consistent with a central role in imprinting. When the *Tsix* knockout comes from the father, embryonic development proceeds normally. Daughters show nonrandom X inactivation with a strong bias toward silencing the knockout chromosome, but they are healthy. This is consistent with earlier work using ES cells. By contrast, most embryos of both sexes die if they inherit the *Tsix* knockout from their mother. This is because the extraembryonic tissues silence the only X in sons and both Xs in daughters. A $Tsix^{-}$ chromosome

‡E-mail: rkelley@bcm.tmc.edu (R. L. K.), mkuroda@bcm.tmc.edu (M. I. K.)

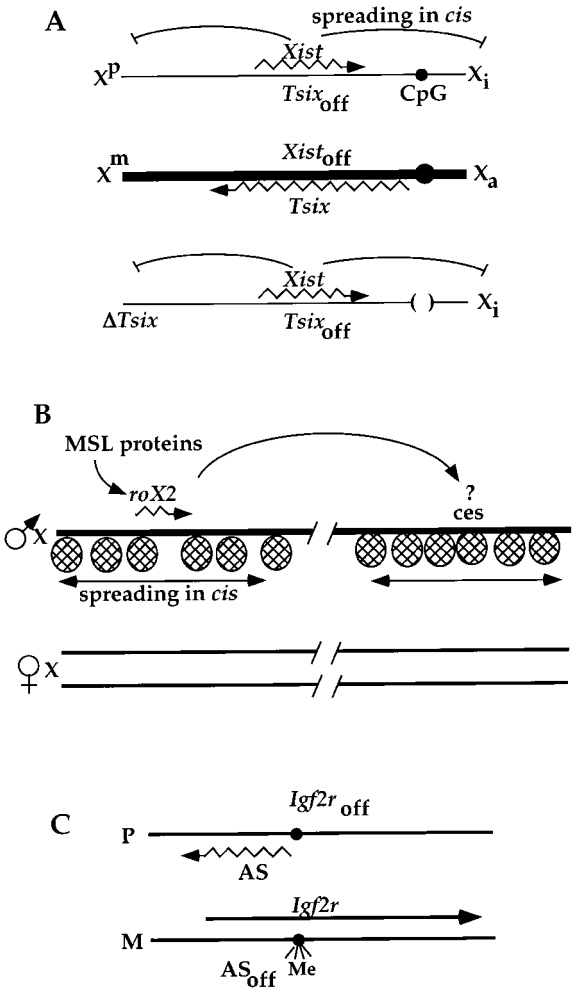

Figure 1. The *Xist* RNA Spreads along the Dosage Compensated X Chromosome like the *Drosophila roX* RNAs, and Shows Imprinting Control by Antisense Transcription as Seen with the *Igf2r* Locus

(A) The murine X inactivation center can be transcribed in the sense direction to make *Xist*, or the antisense direction to make *Tsix*. *Xist* RNA somehow spreads in *cis* along the length of the chromosome to initiate silencing (X_i symbolized by a thin line). *Tsix* negatively regulates *Xist* in *cis* (X_a symbolized by a heavy line; M, maternal). Imprinting may depend upon CpG sequences (filled dot) near the *Tsix* promoter. Blocking *Tsix* transcription allows *Xist* expression to always silence the paternal X chromosome (P) in extraembryonic tissues. This imprint is erased in the cells giving rise to the embryo that must then count the number of X chromosomes and randomly choose one to remain active. Oocytes carrying a deletion of the *Tsix* promoter are unable to mark the maternal Xic and thus their X chromosomes behave as if they are paternally imprinted.

(B) The *roX* RNAs in flies are synthesized from the X and are captured by MSL proteins. This produces a large chromatin remodeling complex (hatched ovals) which can either spread in *cis* or move to other chromatin entry sites (ces) along the X. The resulting histone H4 acetylation leads to elevated transcription of the male X (thick line) to match the output of the two female X chromosomes (thin lines).

(C) The *Igf2r* gene is silent on the paternal chromosome (P) where antisense transcription occurs. Methylation of an internal CpG-rich region blocks antisense transcription allowing *Igf2r* expression to proceed on the maternal homolog (M).

behaves as a paternally imprinted X regardless of which parent donates it. *Tsix* expression is therefore the likely *cis*-acting imprinting factor that protects the maternal X from silencing.

How could an antisense RNA provide a *cis*-acting chromosomal imprint? *Tsix* transcription might result in nonfunctional dsRNA hybrids with *Xist*. However, although nascent *Tsix* transcripts can be detected by in situ hybridization, it is not known whether *Tsix* function requires *Tsix* RNA. For example, RNA polymerases traversing the chromatin in the antisense direction might somehow block sense *Xist* transcription. Alternatively, the *Xist* and *Tsix* promoters might compete for a limiting enhancer. What would happen to imprinting if the *Tsix* promoter were simply inverted to transcribe away from *Xist*?

For a number of years *Xist* RNA stood as the sole example of a noncoding RNA which seemed to spread over flanking chromatin from its site of synthesis to somehow affect chromosome structure. However, the recent discovery of the *roX* RNAs in flies provided a second example of this new class of RNAs and highlights intriguing mechanistic differences between flies and mammals.

roX *RNAs in Flies*

Dosage compensation in fruitflies clearly operates by a different mechanism from mammalian X inactivation. Female flies actively transcribe both X chromosomes at a basal rate. In contrast, males double that rate of transcription for most genes along the single X to synthesize the same amount of products as made in females. Six protein components have been identified which form a complex that mediates this male-specific hypertranscription. One protein, MOF, is a histone H4 acetyltransferase (Smith et al., 2000; Akhtar and Becker, 2000) and another, JIL-1, can phosphorylate histone H3 in vitro (Jin et al., 1999), supporting the idea that dosage compensation is mediated by changes in chromatin architecture. Mutations in five of these genes have been isolated. They prevent hypertranscription of the X resulting in a male-specific lethal phenotype, and so are known collectively as the MSL proteins. The MSL complex has been most extensively studied in cells with large polytene chromosomes where it is localized almost exclusively to hundreds of bands along the male X.

The MLE protein was the first of the MSLs to be molecularly characterized. Its predicted sequence resembles RNA helicases providing an early hint that dosage compensation in flies might involve an RNA component. More recently, two other MSL subunits were shown to bind RNA through their chromodomains (Akhtar et al., 2000). A conceptual breakthrough came when the MSL complex was found to contain at least two RNAs, *roX1* and *roX2* (*R*NA *o*n the *X*), which are 3.7 and 0.6 kb, respectively. These are spliced RNAs with almost no primary sequence homology. The most distinct feature of *roX* RNAs is that they both paint the length of the male X in a finely banded pattern (Franke and Baker, 1999; Meller et al., 2000). Like *Xist* in mammals, each *roX* gene maps to the X chromosome, and the RNAs appear to spread in *cis* from their sites of synthesis when complexed with the MSL proteins (Figure 1B). This proposal is based on the observation that when a *roX* transgene is moved to an autosome, the MSL/*roX* complex is sometimes found in a local region a few hundred kilobases on either side of the insertion site. The affected chromatin suffers ectopic histone H4 acetylation, suggesting that the autosomal genes which were never before targets of dosage compensation may now be hyper-

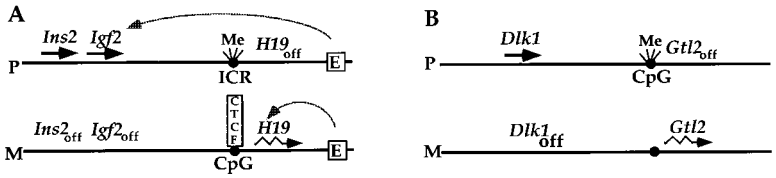

Figure 2. Reciprocal Imprinting of Adjacent Protein-Coding and Noncoding RNA Genes

(A) Reciprocal imprinting is observed at the *Igf2* locus where a single enhancer drives expression of either *Ins2* and *Igf2* from the paternal homolog, or noncoding *H19* RNA from the maternal chromosome. The imprinting control region (ICR) contains CpG-rich sequences that serve as the binding site for the chromatin insulator CTCF. When CTCF occupies the ICR, the enhancer cannot act on distal genes, and *H19* is transcribed. Paternal methylation of the ICR prevents CTCF binding allowing the enhancer to act on the *Igf2* and *Ins2* genes (Hark et al., 2000; Bell and Felsenfeld, 2000).

(B) The paternally imprinted *Dlk1* gene is upstream of a maternally imprinted *Gtl2* gene, which makes a noncoding RNA. As with the *Igf2*/*H19* locus, a differentially methylated CpG-rich region (filled circle) lies upstream on the noncoding RNA gene.

transcribed (Kelley et al., 1999). We have proposed that these unusual RNAs reach their targets by a poorly understood spreading mechanism that is largely independent of the sequence of the affected genes.

The MSL complex appears to initiate spreading at ~35 chromatin entry sites scattered along the length of the X. These sites were first noticed as the probable locations where the individual MSL proteins assemble into functional complexes. The *roX* genes are located at two chromatin entry sites, leading to a model in which MSL proteins capture nascent *roX* transcripts and package them into mature complexes able to spread into flanking chromatin (Kelley et al., 1999) (Figure 1B). So far no additional *roX* RNAs have been identified and the nature of the other chromatin entry sites remains a mystery. While *roX* RNAs have been postulated to spread at most about 1 Mbp from each of 35 chromatin entry sites on the fly X chromosome, *Xist* RNA appears to spread >100 Mbp from its site of synthesis on the mammalian X.

So far *roX* and *Xist* are the only known RNAs thought to spread along a chromosome from their sites of synthesis, but there is no hint of a common evolutionary origin. Dosage compensation arose independently in each animal long after they shared a common ancestor. Moreover, the details of dosage compensation in flies and placental mammals are quite different. *roX* RNAs act with the MSL proteins in males to hypertranscribe the single X. *Xist* acts in females to silence one of two X chromosomes. *Xist* RNA has not been observed to diffuse from its site of synthesis and later reattach to another chromosome. The risk of lethally silencing both female X chromosomes provides a strong selection against such an event. However, in flies, both growing and complete MSL/*roX* complexes appear to exchange freely between the various chromatin entry sites in a soluble form (Kelley et al., 1999; Meller et al., 2000).

If *Xist* is regulated by its antisense *Tsix* gene, could anything analogous occur in flies? In terms of imprinting, the answer is no. Males usually inherit their X chromosome from their mother, but genetic tricks can be used to generate sons carrying a paternal X. In either case dosage compensation occurs in all sons. Flies also have little or no DNA methylation, which seems to play a central role in imprinting in mammals. However, little is known about sex-specific *roX* RNA regulation. Although *roX* RNAs are dependent upon MSL proteins for stability and therefore cannot survive if ectopically expressed in females (Meller et al., 2000), it is not known whether their synthesis is also regulated. Therefore, possible roles for antisense *roX* transcription have not been ruled out.

Beyond the X: Imprinting

Alleles silenced via imprinting are not thought to be packaged in densely compacted heterochromatin, but several imprinted genes have oppositely imprinted noncoding RNA partners nearby on the chromosome. Maternal expression of the *Igf2r* gene is reminiscent of the *Xist*/*Tsix* locus in that imprinting is tied to antisense transcription on the paternal chromosome (Wutz et al., 1997) (Figure 1C). Perhaps the most intensively studied example of imprinting is the paternally expressed *Ins2 Igf2* gene cluster that is linked to the maternally expressed *H19* gene (Tilghman, 1999) (Figure 2A). Although no function has been found for the noncoding *H19* RNA, it is conserved and abundantly expressed. Moreover, the *Dlk1*/*Gtl2* locus was recently found to have a nearly identical arrangement of a paternally imprinted protein coding gene adjacent to a maternally imprinted noncoding RNA, suggesting some type of selection for these unusual RNAs (Schmidt et al., 2000) (Figure 2B).

As unconventional RNAs are being encountered in novel epigenetic regulatory mechanisms, one striking feature is that the site of synthesis is critical to function. In both mammals and flies, moving the *Xist* or *roX* genes from the X to autosomes redirects dosage compensation to the new sites of insertion. Antisense *Tsix* expression from one homolog cannot regulate the *Xist* allele from the other homolog, demonstrating exclusive *cis* activity. The same appears to be true within the *Igf2r* locus. Is it possible that RNA is a common epigenetic regulator? By analogy to the RNAs discovered so far, noncoding RNAs could function by repackaging a local segment of chromatin, or by capturing a complementary mRNA, or they could simply be by-products from some type of mutually exclusive action of linked promoters.

Selected Reading

Akhtar, A., and Becker, P.B. (2000). Mol. Cell *5*, 367–375.

Akhtar, A., Zink, D., and Becker, P. (2000). Nature *407*, 405–409.

Bell, A.C., and Felsenfeld, G. (2000). Nature *405*, 482–485.

Franke, A., and Baker, B.S. (1999). Mol. Cell *4*, 117–122.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). Nature *405*, 486–489.

Jin, Y., Wang, Y., Walker, D.L., Dong, H., Conley, C., Johansen, J., and Johansen, K.M. (1999). Mol. Cell *4*, 129–135.

Kelley, R.L., Meller, V.H., Gordadze, P.R., Roman, G., Davis, R.L., and Kuroda, M.I. (1999). Cell *98*, 513–522.

Lee, J.T. (2000). Cell *103*, this issue, 17–27.

Lee, J.T., and Lu, N. (1999). Cell *99*, 47–57.

Meller, V.H. (2000). Trends Cell Biol. *10*, 54–59.

Meller, V.H., Gordadze, P.R., Park, Y., Chu, X., Stuckenholz, C., Kelley, R.L., and Kuroda, M.I. (2000). Curr. Biol. *10*, 136–143.

Schmidt, J.V., Matteson, P.G., Jones, B.K., Guan, X.-J., and Tilghman, S.M. (2000). Genes Dev. *14*, 1997–2002.

Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). Mol. Cell. Biol. *20*, 312–318.

Solter, D., and Wei, G. (1997). Genes Dev. *11*, 153–155.

Tilghman, S.M. (1999). Cell *96*, 185–193.

Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., and Barlow, D.P. (1997). Nature *389*, 745–749.